Chapter 13

Capture and Identification of miRNA Targets by Biotin Pulldown and RNA-seq

Shen Mynn Tan and Judy Lieberman

Abstract

MicroRNAs (miRNAs) are small noncoding RNAs that regulate the stability and expression of target RNAs in a sequence-dependent manner. Identifying miRNA-regulated genes is key to understanding miRNA function. Here, we describe an unbiased biochemical pulldown method to identify with high-specificity miRNA targets. Regulated transcripts are enriched in streptavidin-captured mRNAs that bind to a transfected biotinylated miRNA mimic. The method is relatively simple, does not involve cross-linking and can be performed with only a million cells. Addition of an on-bead RNase digestion step also identifies miRNA recognition elements (MRE).

Key words MicroRNA, Noncoding RNA, miRNA recognition element, Target identification

1 Introduction

More than a thousand miRNAs are expressed in human cells, each of which can regulate hundreds of coding and noncoding RNAs. miRNAs regulate as many as 90 % of mRNAs [1–3]. miRNAs recognize the genes they regulate by base pairing. Because the miRNA sequence is short (~22 nt) and is not perfectly complementary to target gene sequences, predicting regulated genes is not straightforward. Sequence properties of miRNA-regulated targets have led to a series of canonical rules that are the basis for algorithms that predict miRNA targets. These rules emphasize the importance of exact base pairing of a sequence in the 3′UTR of the regulated transcript to the seed region (nucleotides 2–9) of the miRNA. However, many miRNA-target RNA interactions do not obey the canonical rules. Thus even the best target-prediction algorithms have poor sensitivity and specificity for predicting regulated genes [4]. Since the transcripts of genes that are regulated by a miRNA are generally down-regulated by miRNA over-expression or up-regulated by miRNA knockdown [4], the predictive power of algorithms can be enhanced by combining lists of predicted

Erik Dassi (ed.), *Post-Transcriptional Gene Regulation*, Methods in Molecular Biology, vol. 1358,
DOI 10.1007/978-1-4939-3067-8_13,

targets with experimental data on changes in gene expression that occur with manipulating miRNA expression in cells of interest.

Biochemical isolation of miRNA-induced silencing complexes using antibodies to AGO or other protein components (miRISC) and cross-linking to fix AGO-associated RNAs with techniques such as HITS-CLIP [5, 6], CLIP-seq [7], PAR-CLIP [8], and CLASH [9] provides a global snapshot of miRNA-RNA interactions. However, these methods require large numbers of cells and are technically challenging to perform. They also tend to be less effective for identifying the targets of poorly expressed miRNAs. Identifying within these data the targets of an individual miRNA often requires assumptions about miRNA recognition elements, which introduces considerable bias. The CLASH method has the advantage that it directly identifies miRNA:target miRNA recognition element (MRE) pairs by ligating them and then sequencing the ligated gene product. A recent study, however, suggests that it may be possible to identify such pairs even without adding a ligation step, because of spontaneous ligation by endogenous ligases during the reaction in cell lysates [10]. The UV cross-linking step central to these methods is inefficient, introduces background, and is known to increase the false discovery rate for genome-wide studies of RNA-protein interactions [11].

Streptavidin beads can be used to capture RNAs bound to a transfected miRNA mimic biotinylated at the 3′-end of the active strand [12–19]. These pulldown procedures can identify the transcripts regulated by a specific miRNA in a specific cell type without bias and with a high degree of specificity (70–90 % in several studies) [12–14]. The transfected miRNA mimic is incorporated into the miRISC and is active. The pulldown procedure isolates miRNA:mRNA pairs associated with the miRISC since AGO depletion largely eliminates recovery of specific target mRNAs [13]. These protocols do not require cross-linking and can be performed with just a million transfected cells. When the abundance of RNAs bound to the biotinylated mimic of interest is compared to those bound to a control biotinylated miRNA mimic by real-time PCR, microarray or RNA-seq analysis, hundreds of potential targets are identified. When they are examined experimentally, the great majority of these are *bona fide*. Here, we detail our optimized version of a pulldown protocol that identifies coding and non-coding RNAs regulated by a miRNA. It is still unclear how much the set of miRNA-regulated genes or the extent of gene regulation of each gene varies in different cell contexts. However, transcripts that might be regulated in one cell type may not be identified by the pulldown in another cell type either because the miRNA is not as abundant, the target is not expressed, or the other expressed transcripts in the cell compete for miRNA binding [20]. We also describe a modification that enables MREs to be identified. Because the number of genes that can in principle be regulated by an individual miRNA is large, figuring out the biological functions of a

miRNA from the lists of potential target genes is difficult. Here we describe a bioinformatics strategy that has enabled us to interpret the biological meaning of pulled down genes to rapidly (with minimal trial and error) identify overall miRNA function in cells.

2 Materials

Unless otherwise indicated, all solutions are prepared with nuclease-free water and nuclease-free grade reagents (Ambion, USA). Store all reagents on ice when in use; we recommend that fresh buffers be used for each experiment, unless indicated otherwise. Diligently follow all waste disposal regulations when disposing of waste materials (*see* **Note 1**).

2.1 Transfection and miRNA Mimics

1. Transfection reagent: DharmaFECT 1 (Dharmacon, USA).
2. 3′ biotinylated miRNA mimics: Control—cel-miR-67, and experimental (e.g., hsa-miR-522), with Dharmacon's proprietary modifications. Biotin is coupled to the 3′-end of the active strand of the miRNA (Dharmacon, USA).

2.2 Bead Preparation and Pulldown

1. Streptavidin magnetic beads: Dynabeads® M-280 (Life Technologies, USA).
2. Magnetic rack: DynaMag™-2 Magnet (Life Technologies, USA).
3. Tube rotator in a cold room.
4. Wash buffer: 5 mM Tris–HCl (pH 7), 500 μM EDTA, 1 M NaCl. Add 75 μl of 1 M Tris–HCl (pH 7), 15 μl of 0.5 M EDTA and 3 ml of 5 M NaCl to 12 ml of nuclease-free water. Store at room temperature.
5. Solution A: 50 mM NaCl, 100 mM NaOH. Add 100 μl of 5 M NaCl and 100 μl of 10 N NaOH to 9.8 ml of nuclease-free water. Store at room temperature.
6. Solution B: 100 nM NaCl. Add 100 μl of 5 M NaCl to 4.9 ml of nuclease-free water. Store at room temperature.
7. Phosphate-buffered saline (PBS).
8. TrypLE (Life Technologies, USA).
9. Lysis buffer: 20 mM Tris–HCl (pH 7), 100 mM KCl, 5 mM $MgCl_2$, 25 mM EDTA, 0.3 % NP-40 (Fluka, USA) with 1× Proteinase Inhibitor cocktail. Add 1 ml of 1 M Tris–HCl (pH 7), 2.5 ml of 2 M KCl, 250 μl of 1 M $MgCl_2$, 2.5 ml of 0.5 M EDTA and 150 μl of NP-40 to 43.6 ml of nuclease-free water. Add 1 tablet of Proteinase Inhibitor cocktail (Roche, USA) and mix gently by inversion till completely dissolved.
10. Blocking buffer: 1 mg/ml Ultrapure BSA, 200 μg/ml Yeast t-RNA in lysis buffer. Add 20 μl of 50 mg/ml Ultrapure BSA

and 20 μl of 10 mg/ml Yeast t-RNA (both from Ambion, USA) to 1 ml of lysis buffer. Scale up as needed.

11. Lysis buffer plus: 2.5 mg/ml Ficoll PM400, 7.5 mg/ml Ficoll PM70, 250 μg/ml Dextran Sulfate 670k, 200 U/ml RNaseOUT™, 100 U/ml SUPERase.In™ in Lysis Buffer. Weigh 2.5 mg of dextran sulfate 670k (Fluka, USA), 25 mg of Ficoll PM400, and 75 mg of Ficoll PM70 (GE Healthcare, USA) in a 50 ml tube, add 10 ml of lysis buffer and mix gently by inversion till completely dissolved. Add 50 μl of 40 U/μl RNaseOUT™ and 50 μl of 20 U/μl SUPERase.In™ (Life Technologies, USA) and mix by inversion. Scale up as needed.
12. RNase lysis buffer: 2.5 mg/ml Ficoll PM400, 7.5 mg/ml Ficoll PM70, 250 μg/ml dextran sulfate 670k, 25,000 U/ml RNase T1 in lysis buffer. Weigh 500 μg of dextran sulfate 670k (Fluka, USA), 5 mg of Ficoll PM400, and 15 mg of Ficoll PM70 (GE Healthcare, USA) in a 50 ml tube, add 2 ml of lysis buffer and mix gently by inversion till completely dissolved. Add 50 μl of 1000 U/μl RNase T1 (Fermentas, USA) and mix by inversion. Scale up as needed.

2.3 RNA Precipitation and Cleanup

1. Trizol LS (Life Technologies, USA).
2. Phenol.
3. Chloroform.
4. GlycoBlue (Ambion, USA).
5. Ethanol.
6. Phenol:chloroform:isoamyl (25:24:1).
7. Chloroform:isoamyl (24:1).
8. 3 M sodium acetate pH 5.5 (Ambion, USA).
9. NucAway spin columns (Ambion, USA).
10. PNK buffer and enzyme (NEB, USA).

2.4 RNA-seq Library Preparation

1. Ribo-Zero rRNA removal kit (Epicentre, USA).
2. NEBNext Ion Torrent mRNA library preparation kit (NEB, USA) or equivalent for the Ion Torrent platform, for target transcript identification.
3. NEBNext Illumina small RNA library preparation kit (NEB, USA) or equivalent for the Illumina platform, for target MRE identification.

2.5 Real-Time PCR

1. High-capacity cDNA archive kit (ABI, USA).
2. Fast SYBR green mastermix (Bio-Rad, USA).

2.6 MRE Validation

1. psiCHECK2 plasmid (Promega, USA).
2. NotI and XhoI restriction enzymes and buffers (NEB, USA).

3. CIP (NEB, USA).
4. Quick ligation kit (NEB, USA).
5. One Shot TOP10 competent cells (Life Technologies, USA).
6. Phusion HF PCR kit (NEB, USA).
7. Dual-luciferase reporter assay system (Promega, USA).

3 Methods

3.1 Transfection

1. Optimize the miRNA transfection protocol for your cell line of interest. Transfect enough cells with 50 μM of biotinylated miRNA mimic, to harvest between 1 and 4 million cells for each miRNA (Control and Experimental) at 50 % confluency for lysis (*see* **Note 2**).
2. For MDA-MB-468 breast cancer cell lines, we transfect 1.5 million cells with DharmaFECT 1 and 50 μM of biotinylated miRNA mimic in a 60 mm tissue culture dish as per the manufacturer's instructions, and harvest the cells for the pull-down experiment 20 h after transfection (*see* **Note 3**).

3.2 Bead Preparation

Perform all steps at room temperature. Wash buffer, solution A, and solution B can be kept for up to a month at room temperature.

1. Resuspend magnetic beads (50 μl per 1.5 ml Eppendorf tube per experiment or sample) in 1 ml of wash buffer (maximum of 300 μl of beads per 1 ml of buffer).
2. Mix by inversion and place tubes on magnetic rack for 1 min. Remove wash buffer and repeat twice (for a total of three times).
3. Resuspend beads in solution A.
4. Mix by inversion and place tubes on magnetic rack for 1 min. Remove solution A and repeat once (for a total of two times).
5. Resuspend beads in solution B.
6. Mix by inversion and place tubes on magnetic rack for 1 min. Remove solution B completely.
7. Resuspend beads in blocking buffer (maximum of an equivalent of 150 μl of beads per 1 ml of buffer).
8. Leave tubes on rotator in 4 °C for a minimum of 2 h.

   Perform the following steps during the 20-min cell lysis incubation step in Subheading 3.3 below
9. Place tubes on magnetic rack for 1 min and remove blocking buffer.
10. Resuspend beads in lysis buffer.
11. Mix by inversion and place on magnetic rack for 1 min. Remove lysis buffer and repeat once (for a total of two times).

12. Mix by inversion and place on magnetic rack for 1 min. Remove lysis buffer completely.
13. Combine all magnetic beads and resuspend in lysis buffer plus (110 µl of lysis buffer plus for each sample).
14. For each sample, aliquot 100 µl of the beads/lysis buffer plus mixture into a fresh 1.5 ml Eppendorf tube and keep on ice.

3.3 Cell Lysis and Pulldown Incubation

1. Wash all plates with 1 ml of ice-cold PBS each, and remove PBS completely.
2. Add 500 µl of TrypLE per plate and place in the incubator for 5–15 min (depending on cell line), until cells detach completely from the tissue culture surface. Pipette up and down as necessary.
3. Neutralize TrypLE by adding 1 ml of medium with serum and transfer to a 1.5 ml Eppendorf tube.
4. Spin tubes at 4 °C at 500 × *g* for 5 min, discard supernatant, and keep cell pellet.
5. Resuspend cell pellet in 1 ml of ice-cold PBS.
6. Spin tubes at 4 °C at 500 × *g* for 5 min, remove PBS, and keep cell pellet. Repeat PBS wash once (for a total of two times), taking care not to touch the cell pellet.
7. Remove PBS completely and resuspend cell pellet in 500 µl of ice-cold lysis buffer plus (*see* **Note 4**).
8. Incubate on ice for 20 min, inverting occasionally. Wash beads that are in blocking buffer during this step (*see* Subheading 3.2).
9. Spin tubes at 4 °C at 5000 × *g* for 5 min to pellet nuclei.
10. Remove post-nuclear supernatant, taking care not to touch the nuclear pellet. For each sample, set aside 50 µl of supernatant (cytoplasmic lysate) at 4 °C for analysis of total cellular RNA.
11. Add 400 µl of supernatant (cytoplasmic lysate) to 100 µl of prepared beads in lysis buffer plus for each tube (as prepared in Subheading 3.2).
12. Make sure that all tubes are sealed, and leave on rotator in 4 °C for 4 h.

3.4 Bead Washing and RNA Precipitation

1. Place bead-containing tubes on magnetic rack for 1 min.
2. Remove cytoplasmic lysate from all tubes and add 1 ml of lysis buffer.
3. Mix by inversion and place on magnetic rack for 1 min. Remove lysis buffer and repeat wash steps four more times (for a total of five times).
4. For the last wash step, mix by inversion and place on magnetic rack for 2 min. Remove lysis buffer completely.

5. For each bead-containing sample, resuspend beads in 100 µl of lysis buffer and transfer to a fresh 1.5 ml eppendorf tube. For the total RNA sample, add 50 µl lysis buffer (to achieve a final volume of 100 µl), and treat these samples like the bead-containing samples in subsequent steps.
6. Add 500 µl of TRIzol LS to each tube and mix vigorously.
7. Incubate at room temperature for 5 min, mixing occasionally.
8. Add 100 µl of chloroform to each tube and mix vigorously.
9. Spin tubes at 4 °C at >16,000 × *g* for 15 min.
10. For each sample, transfer the aqueous phase (around 250 µl) to a fresh 1.5 ml Eppendorf tube, taking care not to touch the interphase.
11. For each tube, add 5 µl of GlycoBlue, and subsequently 850 µl of 100 % ethanol.
12. Mix vigorously and precipitate overnight at −20 °C.

3.5 RNA Cleanup

1. The next day, spin tubes at 4 °C at >16,000 × *g* for 30 min. The blue pellet contains the pulldown RNA. Remove ethanol without touching the pellet.
2. For each sample, add 1 ml of 75 % ethanol and mix by inversion.
3. Spin tubes at 4 °C at >16,000 × *g* for 15 min. Remove ethanol without touching the pellet and repeat 75 % ethanol wash step once (for a total of two times).
4. Remove ethanol and dry blue RNA pellets in an open fume hood, taking care not to over-dry them as this interferes with resuspension.
5. Resuspend RNA pellets in 200 µl of nuclease-free water and place on ice.
6. For each tube, add 200 µl of phenol:chloroform:isoamyl (25:24:1).
7. Mix vigorously and spin at >16,000 × *g* for 5 min at room temperature.
8. Transfer the aqueous phase (around 200 µl) to a fresh 1.5 ml Eppendorf tube, taking care not to touch the interphase.
9. For each tube, add 200 µl of chloroform:isoamyl (24:1).
10. Mix vigorously and spin at >16,000 × *g* for 5 min at room temperature.
11. Transfer the aqueous phase (around 200 µl) to a fresh 1.5 ml eppendorf tube, taking care not to touch the interphase.
12. For each tube, add 20 µl of 3 M sodium acetate pH 5.5, 4 µl of GlycoBlue, and subsequently 1 ml of 100 % ethanol.
13. Mix vigorously and precipitate overnight at −80 °C.

14. The next day, spin tubes at 4 °C at >16,000 × *g* for 30 min. The blue pellet contains the RNA. Remove ethanol without touching the pellet.
15. For each sample, add 1 ml of 75 % ethanol and mix by inversion.
16. Spin tubes at 4 °C at >16,000 × *g* for 15 min. Remove ethanol without touching the pellet and repeat 75 % ethanol wash step once (for a total of two times).
17. Remove ethanol and dry blue RNA pellets in open fume hood, taking care not to over-dry them as this interferes with resuspension.
18. Resuspend RNA pellets in 50 μl of nuclease-free water and place on ice. The pulldown RNA is now ready for downstream analyses (*see* **Note 5**).

3.6 Pulldown Analysis

1. The pulled down RNA can be analyzed by microarray or RNA-seq. This chapter provides the details for RNA-seq (for a microarray protocol, *see* [12]). RNA-seq has the advantage that it provides information for bound RNAs of all classes including ncRNAs without being limited to particular splice variants of mRNAs probed on the array. It can also distinguish genes from highly homologous pseudogenes and alternately spliced or polyadenylated transcripts. However, library preparation is more complicated, since very abundant rRNAs need to be depleted or they will dominate the reads. Moreover the analysis of RNA-seq data is more difficult than microarray data. Essentially for either method, the fold enrichment of the miRNA targets is calculated as the normalized ratio of the experimental pulldown signal to control pulldown signal (*see* **Note 6**). For microarray data, normalization of the ratio of the experimental vs. control pulldown to the ratio of the experimental vs control input samples greatly improved the identification of bona fide targets [12]. This improvement may have been for two reasons. First, background binding of mRNAs is enhanced when they are more abundant and normalization to the input takes this bias into account. Second, transfection of the miRNA mimic reduces the amount of target mRNA. Therefore, for most targets, whose mRNA is reduced by transfection of the miRNA, the normalized enrichment ratio (experimental PD/control PD)/(experimental input/control input) goes up because the numerator increases, while the denominator decreases. Biological replicates or triplicates are required to achieve statistically meaningful results. The choice of a cutoff to separate targets from other genes is in some ways arbitrary. We use a fold enrichment cutoff of 2 and p-value <0.01, which results in a target gene list of a few hundred RNAs (*see* **Note 7**). To identify miRNA targets using microarrays, we first normalize the data (platform dependent; for

Illumina Beadchips, we use cubic spline without background normalization), and use a differential gene expression approach that compares the normalized signal of control and experimental pulldowns for each gene to calculate fold enrichment and *p*-values. This can be performed using the NIA Array Analysis software available online (http://lgsun.grc.nia.nih.gov/ANOVA/index.html) [13].

2. For RNA-seq, the coverage required for miRNA target identification can be lower than transcriptome-wide differential expression RNA-seq analyses, if we assume that the number of targets for a miRNA that we can detect is in the thousands at most (i.e., at most a few percent of the transcriptome). We recommend at least biological replicates, with more than 200 million uniquely-mapped bases per replicate. This should be achievable with most RNA-seq platforms, for example Ion Torrent from Life Technologies, or even with multiplexing using higher throughput equipment (for example the HiSeq platforms from Illumina).

3.7 Pulldown-seq Library Prep and Analysis

For Pulldown-seq library preparations, the use of straightforward kits and rRNA removal are strongly recommended. For Pulldown-seq on the Ion Torrent platform, we performed Ribo-Zero rRNA removal before preparing the library with the NEBNext mRNA library preparation kit with Ion Torrent-specific primers, following the manufacturer's protocol.

The bioinformatics analysis workflow for Pulldown-seq on the Ion Torrent platform as previously published is as follows [13]:

1. Screen pre-alignment and post-alignment libraries for quality, specificity of mapping and containment sequences using FASTQC, RSeQC and RNA-SeQC.
2. Trim low quality bases, homopolymer sequences and sequences matching the first 13 bases, and the reverse complement of the adapter sequences for Ion Torrent using cutadapt, and discard trimmed reads smaller than 30 nt.
3. Align trimmed reads to the latest build of the human genome using Novoalign (with the parameters: -H -k -n 250 -F STDFQ -r all 10 -e 10 -g 15 -x 4), keeping only uniquely mapped reads.
4. Generate post-alignment gene counts using htseq-count, with the counts aggregated by gene_id.
5. Call differentially expressed (or in this case differentially enriched) transcripts using the default options of DESeq, with the fold enrichment and *p*-value cutoffs at 2 and 0.05, respectively.

3.8 Target Validation

1. The first step to validate miRNA targets is to verify the microarray or RNA-seq enrichment in the pulldown by real-time PCR quantification of the RNA in the miRNA and control pulldown relative to its level in total cellular RNA, normalized to a house-

keeping gene mRNA such as *GAPDH*. To do this we synthesize cDNA from 2 μl of pulldown RNA from both control and experimental miRNA pulldowns in a total reaction volume of 30 μl, and perform SYBR green real-time PCR on tenfold diluted cDNA with custom-designed primers for a set of 10 or more target genes. These are selected from the microarray or RNA-seq results to represent the spectrum of fold enrichment and *p*-values. The real-time PCR fold enrichment values are then compared to the microarray or RNA-seq results for each specific target to obtain the R-squared value (*see* **Notes 5**, **8**, and **9**).

2. Next we determine whether transfection of the miRNA decreases the target gene mRNA and protein. For mRNA expression, we perform real-time PCR with cDNA synthesized from total RNA from both control and experimental miRNA transfected cells and analyze the reduction of expression for each specific target (*see* **Note 10**). The proportion of all the pulldown-predicted miRNA target RNAs that are downregulated can also be analyzed using microarrays [12–14] or RNA-seq. For protein expression, western blots of cell extracts, comparing cells transfected with control and experimental miRNAs, are probed with antibodies to proteins of interest [12–14]. Because western blots are not always quantitative, it is a good idea to compare protein levels using serial dilutions of samples, with a range of cell equivalents in each lane. The optimal time to examine target gene expression may be later than what was optimized for the pulldown, since more time is needed to downregulate gene expression. For most cells and miRNAs, testing gene regulation should be performed 48–72 h after transfection, with the later time point preferred for some genes.

3.9 RNase Treatment for Identification of MREs

The addition of RNase treatment to the above pulldown protocol will enable the identification of MREs via RNA-seq, termed IMPACT-seq for *i*dentification of *M*REs by *p*ulldown and *a*lignment of *c*aptive *t*ranscript-*seq*uencing [13]. If this is desired, please perform all steps from Subheadings 3.1–3.6, but replace Subheading 3.5 with this section. As this protocol introduces additional wash steps and removal of RNAs below 20 nt with NucAway spin columns, we generally combine the RNA for three samples for each experiment per RNA-seq reaction.

1. Remove cytoplasmic lysate from all tubes and add 1 ml of lysis buffer.
2. Mix by inversion and place on magnetic rack for 1 min. Remove lysis buffer and repeat wash steps four more times (for a total of five times). Remove lysis buffer for the final wash completely.
3. For each bead-containing sample, resuspend beads in 400 μl of ice-cold RNase lysis buffer, transfer to a fresh 1.5 ml Eppendorf tube and place on ice.

4. When all tubes are ready, incubate them in a 37 °C water bath or heat block for 10 min with occasional inversions to ensure an even temperature in each tube.
5. Place all tubes on ice for 5 min.
6. Place tubes on magnetic rack for 1 min, and remove RNase lysis buffer completely.
7. Add 1 ml of lysis buffer for each tube, mix by inversion and place on magnetic rack for 1 min. Remove lysis buffer and repeat wash steps four more times (for a total of five times).
8. For the last wash step, mix by inversion and place on magnetic rack for 2 min. Remove lysis buffer completely.
9. For each set of bead-containing samples, resuspend combined beads (usually three tubes) in 100 μl of lysis buffer and transfer to a fresh 1.5 ml Eppendorf tube.
10. Add 500 μl of TRIzol LS to each tube and mix vigorously.
11. Incubate at room temperature for 5 min, mixing occasionally.
12. Add 100 μl of chloroform to each tube and mix vigorously.
13. Spin tubes at 4 °C at >16,000 × *g* for 15 min.
14. For each sample, transfer the aqueous phase (around 250 μl) to a fresh 1.5 ml Eppendorf tube, taking care not to touch the interphase.
15. For each tube, add 5 μl of GlycoBlue, and subsequently 850 μl of 100 % ethanol.
16. Mix vigorously and precipitate overnight at −20 °C.
17. The next day, spin tubes at 4 °C at >16,000 × *g* for 30 min. The blue pellet contains the MRE RNA. Remove ethanol without touching the pellet.
18. For each sample, add 1 ml of 75 % ethanol and mix by inversion.
19. Spin tubes at 4 °C at >16,000 × *g* for 15 min. Remove ethanol without touching the pellet and repeat 75 % ethanol wash step once (for a total of two times).
20. Remove ethanol and dry blue RNA pellets in an open fume hood, taking care not to over-dry them as this interferes with resuspension.
21. Resuspend RNA pellets in 50 μl of nuclease-free water and place on ice.
22. Tap the NucAway columns to settle dry gel onto the bottom of the spin column.
23. Hydrate the columns with 650 μl of nuclease-free water, cap, vortex, and tap out the air bubbles. Leave the columns at room temperature for 10 min.
24. Place each column in a 2 ml collection tube and centrifuge the column at 700 × *g* for 2 min.

25. Discard the collection tubes, and place each column in a fresh 1.5 ml Eppendorf tube.
26. Add 50 μl of RNA for each sample onto the hydrated gel in each column slowly, and spin immediately at 700 × *g* for 2 min.
27. Discard the columns and top up size-selected RNA to 174 μl with nuclease-free water.
28. To each tube of 174 μl of size-selected RNA, add 20 μl of PNK buffer and 4 μl of PNK enzyme, as well as 1 μl each of RNase OUT Superase IN.
29. Incubate at 37 °C for 1 h.
30. Add 200 μl of phenol:chloroform:isoamyl (25:24:1) to each tube.
31. Mix vigorously and spin at >16,000 × *g* for 5 min at room temperature.
32. Transfer the aqueous phase (around 200 μl) to a fresh 1.5 ml Eppendorf tube, taking care not to touch the interphase.
33. For each tube, add 200 μl of chloroform:isoamyl (24:1).
34. Mix vigorously and spin at >16,000 × *g* for 5 min at room temperature.
35. Transfer the aqueous phase (around 200 μl) to a fresh 1.5 ml Eppendorf tube, taking care not to touch the interphase.
36. For each tube, add 20 μl of 3 M sodium acetate pH 5.5, 4 μl of GlycoBlue, and subsequently 1 ml of 100 % ethanol.
37. Mix vigorously and precipitate overnight at −80 °C.
38. The next day, spin tubes at 4 °C at >16,000 × *g* for 30 min. The blue pellet contains the RNA. Remove ethanol without touching the pellet.
39. For each sample, add 1 ml of 75 % ethanol and mix by inversion.
40. Spin tubes at 4 °C at >16,000 × *g* for 15 min. Remove ethanol without touching the pellet and repeat 75 % ethanol wash step once (for a total of two times).
41. Remove ethanol and dry blue RNA pellets in open fume hood, taking care not to over-dry them as this interferes with resuspension.
42. Resuspend RNA pellets in 20 μl of nuclease-free water and place on ice. The MRE RNA is now ready for RNA-seq.

3.10 IMPACT-seq Library Prep and Analysis

For IMPACT-seq library preparations, the use of small RNA kits is strongly recommended, and rRNA removal is discouraged. For IMPACT-seq on the Illumina platform, we prepared the library with the NEBNext small RNA library preparation kit with Illumina-specific primers, following the manufacturer's protocol. At the

size-selection step of the protocol that is generally performed with gel extraction, we select for an insert size of between 20 and 60 nt.

The bioinformatics analysis workflow for IMPACT-seq on the Illumina platform as previously published is as follows [13]:

1. Screen pre-alignment and post-alignment libraries for quality, specificity of mapping and containment sequences using FASTQC, RSeQC, and RNA-SeQC.
2. Trim low-quality bases, homopolymer sequences, and sequences matching the first 13 bases, and the reverse complement of the adapter sequences for Illumina using cutadapt, and discard trimmed reads smaller than 20 nt.
3. Align trimmed reads to the latest build of the human genome with Tophat, keeping only uniquely mapped reads with ≦2 mismatches.
4. These steps can be performed using the bipy (http://github.com/roryk/bipy) and bcbio-nextgen (http://github.com/chapmanb/bcbio-nextgen) automated sequencing analysis pipelines.
5. Call peaks for both control and experimental miRNA IMPACT-seq samples using CLIPper with the following parameters: --poisson-cutoff=0.05 --superlocal --max_gap=0 --processors=8 -b $file -s hg19 -o $file.
6. Identify MREs by requiring that each peak have ≥5 reads in the experimental miRNA sample with at least twice as many normalized reads in the experimental miRNA sample as control.

3.11 Analysis and Validation of MREs

The analysis of putative IMPACT-seq MREs can be accomplished by cloning a random set of these sequences into psiCHECK2, and performing 3′UTR luciferase assays [12–14].

1. For each IMPACT-seq MRE sequence, synthesize a phosphorylated, double-stranded DNA fragment, with a 5′ XhoI and a 3′ NotI restriction enzyme overhang sequence.
2. Digest psiCHECK2 with XhoI and NotI, and dephosphorylate cut vector with CIP.
3. Ligate dsDNA in **step 1** with cut vector in **step 2**, and transform into competent cells.
4. Pick colonies and perform PCR screening for positive clones with the Phusion kit, using a universal forward primer targeting the Renilla luciferase (ACCCTGGGTTCTTTTCCAAC), and a specific reverse primer that corresponds to the 3′–5′ sequence of the MRE.
5. The positive clones can then be sequenced using the universal forward primer to confirm the exact sequence of MRE that is cloned.

6. A positive control MRE (the reverse complement of the experimental miRNA sequence) and a few negative control MREs (a random sequence or known sequences that are not targets of the experimental miRNA) can also be cloned into psiCHECK2 in this way.
7. These psiCHECK2-MRE vectors are then transfected into the cell line of interest in 24-well plates, with either the control miRNA or the experimental miRNA using the protocol optimized above in Subheading 3.1.
8. At 48 h after transfection, luciferase activity is assayed with the Dual Luciferase Reporter kit and a plate reader.
9. The Renilla luciferase value (where the MRE is cloned) is normalized by the Firefly luciferase value, and this is calculated for both the control and experimental miRNA transfections. The ratio of experimental versus control value for each MRE is then compared to both negative and positive MRE control ratios to determine if the putative IMPACT-seq MRE is functional.

If more stringent validation is required, MRE sequences with a few point mutations of the expected miRNA-binding residues can be cloned into psiCHECK2 and assayed as above, with the expectation that these mutations will disrupt miRNA binding.

3.12 Bioinformatics Analysis of Target Genes

1. miRNAs regulate biological processes by targeting multiple genes involved in similar pathways [4]. To decipher biological functions of a miRNA, targets identified by the pulldown method can be analyzed by gene ontology, interactome and/or pathway analysis software, for example Ingenuity Pathway Analysis (IPA, www.ingenuity.com) and Database for Annotation, Visualization and Integrated Discovery (DAVID, david.abcc.ncifcrf.gov) [12–14].
2. Gene ontology, interactome, and pathway analysis can also be conducted on genes that are downregulated after overexpression of the miRNA. This analysis takes into account genes regulated both directly and indirectly by the miRNA.
3. For pathway analysis of both pulldown targets and downregulated genes, we recommend the use of IPA (www.ingenuity.com). For the two gene lists, first connect all genes that are directly related with the default settings, as defined by the curated IPA database, without selecting the optional predicted miRNA interactions. Second, conduct a Core analysis with the default IPA settings, to generate IPA scores for the top associated network functions and significantly enriched molecular functions for both lists. Finally, identify network and molecular functions that are common in both pulldown targets and downregulated genes. This analysis can be done using different stringency cutoffs to select the target genes. The most important functions will persist with more stringent cutoffs since the

most highly regulated genes are generally most enriched in the pulldown.

4. A useful addition to gene ontology and pathway analyses is to search for over-represented transcription factors predicted to bind to the promoter regions upstream of target genes. This analysis, which can be done using TRANSFAC, a well-curated knowledge base of eukaryotic transcription factors (www.gene-regulation.com), can shed light on miRNA function based on the known functions of the transcription factors that transcriptionally regulate the miRNA target genes [13]. Combining all these methods should result in a set of hypotheses about the set of biological functions of the mRNA that can direct experimental efforts into fruitful investigations that avoid unproductive searching. This streamlined approach should be useful for uncovering the hidden meaning of large genome-wide datasets more generally.

4 Notes

1. We recommend the use of miRNA mimics from Dharmacon, which are biotinylated at the 3′ end of the active strand, and are designed with proprietary modifications to both the active and passive strands to improve loading into the miRISC. Trials with siRNA-like biotinylated dsRNA with and without bulges did not work as well.
2. Optimizing miRNA transfection efficiency for the cell line of interest is crucial for the success of this protocol. This may require testing other transfection lipids or using nucleoporation to transfect cells that are ordinarily more difficult to transfect (such as nonadherent cells, especially hematopoietic cells and lymphocytes). We strongly recommend optimizing transfection to maximize both the number of transfected miRNAs per cell and the uniformity of transfection. One way to test the efficiency of any transfection protocol is to monitor mRNA levels of known miRNA targets 24 or 48 h after transfection via real-time PCR. For miRNAs without any previously validated targets, one approach is to test the top conserved or highest scoring genes predicted by algorithms (for example TargetScan) as positive controls. Another approach is to test the effect of the transfected miRNA on luciferase activity in cells transiently transfected with a 3′UTR dual-luciferase reporter, such as psiCHECK2 [12–14], containing an antisense sequence to the miRNA active strand. Transfection conditions that cause >90 % reduction in expression of known targets or luciferase activity are ideal. Alternatively, one may optimize transfection of fluorescently labeled miRNAs (or siRNAs) using flow cytometry analysis of the miRNA tag to measure miRNA internalization. Low transfection may impair the ability to detect pulled down targets (*see* **Note 11**). In addi-

tion, we recommend performing pulldown experiments in cell lines with a moderate level of expression of your miRNA of interest. This ensures that the transfected biotinylated miRNA mimics will not be diluted out by high levels of the endogeneous miRNA, enabling efficient capture of targets. Performing pulldowns in such cell lines will also allow for identification of biologically relevant targets, as both the miRNA of interest and its targets are expressed in the same system.

3. The length of time between transfection and cell lysis also needs to be optimized for each cell line and miRNA. The optimal timing balances the opposing considerations of increasing binding of the miRNA to miRISC and target RNAs with the miRNA-mediated accelerated degradation of the target mRNA. In our experience, overnight or 20 h is usually ideal. However, for some cells or miRNAs the optimal time, determined using the systems described in **Note 2**, needs to be shortened to optimize the pulldown. To optimize the timing, the fold enrichment of known target mRNAs should be assessed by real-time PCR after pulldowns at 4, 8, 12, 16, and 20 h.
4. It is important to resuspend cells to a single-cell state as much as possible before the addition of Lysis Buffer Plus for efficient lysis.
5. As TRIzol LS is used to precipitate RNA, be cautious about the carry-over of guanidine salts to the RNA prep. To minimize the effect of these salts (which can inhibit PCR reactions), we use only a small amount of RNA (2 μl of the total 50 μl of RNA) for each 30 μl cDNA reaction.
6. Performing the pulldowns with multiple miRNAs in parallel can be useful as these additional miRNAs can be used to optimize the transfection conditions and they can also serve as controls for specificity. Target genes that are enriched for binding to multiple miRNAs may bind nonspecifically.
7. The cutoffs used to define the list of miRNA-regulated genes are in some ways arbitrary. It is unknown in any context how many of the genes that a miRNA can potentially regulate are physiologically important for miRNA function. It is important to keep in mind that the pulldown identifies potentially regulated genes that are isolated under conditions of miRNA overexpression. We have found that the fold enrichment in the pulldown correlates with the extent of gene downregulation by the miRNA [12–14]. Since the most highly enriched mRNAs may be those that are most strongly regulated by the miRNA, using more stringent cutoffs can reduce the gene lists to a more manageable size and may focus further experimentation on genes most likely to be important, however, at the risk of losing important targets. The overlap of the set of genes down-regulated after miRNA overexpression from microarray or RNA-seq data with the set of genes enriched in the pull-

down can also be used to define a more focused candidate target gene set, testing a variety of cutoffs for each component to produce different sized gene lists. One way to determine the best cutoffs is to compare plots of cumulative changes induced by the miRNA in mRNA levels of the gene sets defined by different cutoffs (with statistical analysis using the Kolmogorov-Smirnov test) [13]. Increases in the stringency of the cutoff that reduce the gene list size, but do not significantly change overall gene downregulation, reduce the size of the gene list, but are unlikely to select better targets.

8. In choosing control genes in the analysis of pulldown or knockdown results with real-time PCR, we recommend that multiple negative control housekeeping genes (such as *GAPDH*, *ACTB* and *SDHA*) be tested. Some miRNAs target one or more of the commonly used housekeeping genes. Genes, whose expression does not vary after transfection of the miRNA, should be chosen for normalizing the PCR results.
9. It is useful to include positive controls for known miRNA targets in real-time PCR analyses, if possible. This is a good indication of the efficiency of each pulldown experiment.
10. Since most miRNA targets identified by enrichment in the pulldown are also downregulated by over-expression of the miRNA [12–14], we recommend validating novel miRNA targets by examining changes in their expression after miRNA over-expression.

Although multiple studies from various groups have identified miRNA targets with this technique [12–19], a recent report suggests that biotinylated miRNAs do not associate with AGO, based on analysis of miR-27 [21]. This finding seems to contradict our finding that AGO depletion removes 70–80 % of target gene mRNAs from the miR-522 pulldown [13]. One possible explanation for their negative result could be their low level of transfection of the miRNA mimic. This reinforces our emphasis on the importance of high transfection efficiency (*see* **Note 2**). However, another potential explanation is that the biotin tag may interfere with miRISC/AGO binding of some biotinylated miRNAs. In particular, we were unable to get the pulldown to work well for miR-21 in some cells. This could be a sequence-specific or even a context-specific problem, since posttranslational modifications of AGO can affect miRNA binding [22].

Acknowledgements

We thank Jingmin Jin and Larry McReynolds for their sequencing expertise, Rory Kirchner, Oliver Hofmann and Winston Hide for their bioinformatics expertise, and members of the Lieberman lab for critical discussions. S.M.T. was supported by the Department of Defense (DOD) Breast Cancer Research Program (BCRP).

References

1. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215–233
2. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15–20
3. Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 126:1203–1217
4. Thomas M, Lieberman J, Lal A (2010) Desperately seeking microRNA targets. Nat Struct Mol Biol 17:1169–1174
5. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460:479–486
6. Loeb GB, Khan AA, Canner D, Hiatt JB, Shendure J, Darnell RB, Leslie CS, Rudensky AY (2012) Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. Mol Cell 48:760–770
7. Zisoulis DG, Lovci MT, Wilbert ML, Hutt KR, Liang TY, Pasquinelli AE, Yeo GW (2010) Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. Nat Struct Mol Biol 17:173–179
8. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M et al (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141:129–141
9. Helwak A, Kudla G, Dudnakova T, Tollervey D (2013) Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 153:654–665
10. Grosswendt S, Filipchyk A, Manzano M, Klironomos F, Schilling M, Herzog M, Gottwein E, Rajewsky N (2014) Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. Mol Cell 54:1042–1054
11. Keene JD, Komisarow JM, Friedersdorf MB (2006) RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. Nat Protoc 1:302–307
12. Lal A, Thomas MP, Altschuler G, Navarro F, O'Day E, Li XL, Concepcion C, Han YC, Thiery J, Rajani DK et al (2011) Capture of microRNA-bound mRNAs identifies the tumor suppressor miR-34a as a regulator of growth factor signaling. PLoS Genet 7:e1002363
13. Tan SM, Kirchner R, Jin J, Hofmann O, McReynolds L, Hide W, Lieberman J (2014) Sequencing and systems analysis of captive target transcripts identifies primate-specific miR-522 as an inducer of mesenchymal transition. Cell Rep 8:1225–1239
14. Perdigao-Henriques R, Petrocca F, Altschuler G, Thomas MP, Le MT, Tan SM et al. miR-200 promotes the mesenchymal to epithelial transition by suppressing multiple members of the Zeb2 and Snail1 transcriptional repressor complexes. Oncogene 2015. Advanced doi: 10.1038/onc.2015.69.online publicaton
15. Orom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 30:460–471
16. Cloonan N, Wani S, Xu Q, Gu J, Lea K, Heater S, Barbacioru C, Steptoe AL, Martin HC, Nourbakhsh E et al (2011) MicroRNAs and their isomiRs function cooperatively to target common biological pathways. Genome Biol 12:R126
17. Kang H, Davis-Dusenbery BN, Nguyen PH, Lal A, Lieberman J, Van Aelst L, Lagna G, Hata A (2012) Bone morphogenetic protein 4 promotes vascular smooth muscle contractility by activating microRNA-21 (miR-21), which down-regulates expression of family of dedicator of cytokinesis (DOCK) proteins. J Biol Chem 287:3976–3986
18. Le MT, Shyh-Chang N, Khaw SL, Chin L, Teh C, Tay J, O'Day E, Korzh V, Yang H, Lal A et al (2011) Conserved regulation of p53 network dosage by microRNA-125b occurs through evolving miRNA-target gene pairs. PLoS Genet 7:e1002242
19. Martin HC, Wani S, Steptoe AL, Krishnan K, Nones K, Nourbakhsh E, Vlassov A, Grimmond SM, Cloonan N (2014) Imperfect centered miRNA binding sites are common and can mediate repression of target mRNAs. Genome Biol 15:R51
20. Tay Y, Rinn J, Pandolfi PP (2014) The multilayered complexity of ceRNA crosstalk and competition. Nature 505:344–352
21. Guo YE, Steitz JA (2014) 3′-Biotin-tagged microRNA-27 does not associate with Argonaute proteins in cells. RNA 20:985–988
22. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14:447–459